Compte rendu du TP4 Michel et Maxime

1.1 Figure 1 et Figure 3 sont identiques parce que dans ce cas, quand ρi prend le minimum, on a ρi+P(wi|x)=P(x), c'est a dire P(wi|x) prend le maximum, et de plus P(wi|x)*P(x)=P(x|wi)*P(wi) avec P(wi)=1/C, alors P(x|wi) prend le maximum, c'est le cas de μ et Γ connue, Donc ces deux discriminateurs sont équivalent, les deux figures sont même. Pour le figure 13, dans les images de P(wi|x), ses frontiers de la partie jaune(possibilité haute) representent la frontière de discriminateur, la summation des partie jaune est la square ensemble. Pour les images de P(x|wi), ça signifie que si x appartient à wi, il est plus possible d'apparaître dans les cercles (jaunes, vertes) concentriques de possibilité. Les images dans Figure 13 sont accordé avec Figure 1 et Figure 3.

1.2 Si on presente le nombre dans la matrice sont Ni,j, alors

P(wi|wj)=P(wi,wj)/P(wj)=Ni,j/(N1,j+N2,j+....+Nc,j)

1.3 On observe que N1,2 devient 0 et N2,1 devient très grand, et la surface est plus proche de class C1 qu'avant. Le raison est a1,2 est très grand, P(w1|w2) a une importante contribution au R, alors, il faut P(w1,w2) être très petit, c'est a dire N1,2 devient très petit. Pour realiser ça, il faut diminuer le point reel de class C2, mais estime comme class C1, donc la surface est plus proche de class C1 pour inclure plus de points réels de class C2. En bref, a1,2 est grand signifie que si on judge x de w2(vrai) à w1(faux), le coût est grand. Donc le discriminateur veut éviter cette faute.

1.4 Quand P(w2)=10P(w1),on voit plus de points reel de class C2，N1,2 + N2,2 + N3,2 devient plus grand. La surface qui est contourné par les discriminateur dans le cas Bayes est plus grand. Car P(w2)=10P(w1) signifie qu'un x est plus possible d'appartenir à w2 à priori.

2.1 La performance du discriminateur de Bayes représenté en nombres sont identique avec la figure 3 qui signifie des coefficients de correlation entre les images de chiffre 'i' et 'j'. Quand le nombre dans figure 2 est plus grand(en trace), le carré de figure 3 est plus noir.

2.2 On voit que R Bayes est constante, parce qu'il est equivalent a le cas de μ et Γ connue, et R des deux autre courbes diminue selon Papp auguemente et ils converge vers le cas R Bayes, c'est le meme cas au TP3, l'un est R, l'autre est tau_g. Et R de Bayes lineaire est diminue avant le cas quadratique, fonctionne meilleur. Mais ma question est que on n'eat pas sure la raison. Peut-etre selon Papp auguemente tau_g sont proche pour les different chiffre et Pgen sont le même, donc Γ sont pr èsque egales, correspond au cas linéaire.

2.3 Quand on veut diminue l'influence sur R, on voit que le coefficient de correlation de N8,8 diminue, c'est a dire que quand on utilise le risque de Bayes, si on veut diminue l'influence des chiffres ressembles, comme on augumente a6,8 pour obtenir N6,8 plus petit, ça veut d'autres coefficients de correlation.

3.1 D'après les coefficients de correlation, la sommation de RN_kappa en trace est plus grande que celle de RN_beta, d'ou la performance de RN_kappa est mieux que celle de RN_beta où dans la matrice du coût il y a un grand nombre 100 dans cette colone. La risque de RN_beta est plus petit, d'après l'histogramme, mais ils sont tous plus petit que celle de RN. Quand la base d'apprentissage augmente, par exemple Papp=10000, les performance tendent vers le meme limitation, ils sont prèsque pareille.

3.2 Quand on change la function coût,(choix-cout=3, dans la matrice des couts, il n'y a pas des grands nombres sauf la dernière colone avec des nombres 10), les performances et les risques des deux RN sont prèsque pareilles. Quand on change le prior, l'endroit où le prior est très peu est moins possible d'avoir un point estimée. Les performances et les risques des deux RN sont prèsque les mêmes, car le prior cache la diffèrence généré par la matrice des coûts.

3.3 Comme on a fait dans le TP , on ajoute un paramètre A pour générer une base où l'amplitude est inconnu. Ici, on propose à ajouter plusieurs paramètre [ai,j],[bi] qui représente la matrice des coûts inconnu et la colone du prior inconnu. Et on détermine une l'intervale où les paramètres peuvent bouger au hasard. D'après ce modèle, on regénére la base d'apprentissage avec un grand Papp qui assure l'accurance. Et on utilise cette base à effectuer l'apprentissage. (on n'est pas sûr est-ce que ça peut marcher ou pas. Il faut une grand nombre de Papp.)